U.S. DEPARTMENT OF COMMERCE
National Technical Information Service

PB-265 304

A Quantitative Evaluation of the Pollutant Standards Index

Environmental Protection Agency, Washington, D C

1976

TECHNICAL REPORT DATA
(Please read Instructions on the reverse before completing)

1. REPORT NO.	2.	3. PB 265 304
EPA-600/J-76-034	JOURNAL ARTICLE	

4. TITLE AND SUBTITLE	5. REPORT DATE
A Quantitative Evaluation of the Pollutant Standards Index	6. PERFORMING ORGANIZATION CODE
7. AUTHOR(S)	8. PERFORMING ORGANIZATION REPORT NO.
Wayne R. Ott and William F. Hunt Jr.	
9. PERFORMING ORGANIZATION NAME AND ADDRESS Office of Monitoring & Technical Support Office of Research & Development U.S. Environmental Protection Agency Washington, DC 20460	10. PROGRAM ELEMENT NO. 1HD621 11. CONTRACT/GRANT NO.
12. SPONSORING AGENCY NAME AND ADDRESS U.S. Environmental Protection Agency 401 M St., S.W. Washington, D.C. 20460	13. TYPE OF REPORT AND PERIOD COVERED 14. SPONSORING AGENCY CODE EPA/600/19

15. SUPPLEMENTARY NOTES
Published in the Journal of the Air Pollution Control Association, Vol. 26, No. 11, November 1976, pp. 1050-1054.

16. ABSTRACT

The U.S. Environmental Protection Agency has recommended a daily air pollution index for use by State and local air pollution control agencies--the Pollutant Standards Index (PSI). The new index makes use of "segmented linear functions" which convert measured concentrations of each air pollutant into a normalized number. The overall index is reported for the pollutant with the maximum subindex value, or the "critical pollutant."

To evaluate the index during its design phases, a computer analysis was undertaken using air quality data from selected cities. This analysis examined the statistical properties of the new index, the time series of index values, and the frequency distribution of descriptor words and critical pollutants. The index was found to be an effective tool for compactly representing a year of air quality data at any air monitoring station. It showed sufficient day-to-day variation to give useful information to the public on daily air quality levels. This paper provides sufficient technical detail for the engineer who wishes to apply the index on a large-scale computer, minicomputer, or hand calculator.

17. KEY WORDS AND DOCUMENT ANALYSIS

a. DESCRIPTORS	b. IDENTIFIERS/OPEN ENDED TERMS	c. COSATI Field/Group
System analysis, Statistical analysis, Air pollution, Environmental engineering	Environmental indices Air pollution indices Computer analysis Mathematics operations research	05 09D 12A 12B 13B

18. DISTRIBUTION STATEMENT	19. SECURITY CLASS *(This Report)*	21. NO. OF PAGES
RELEASE TO PUBLIC	UNCLASSIFIED	7
	20. SECURITY CLASS *(This page)*	22. PRICE PC A02 MF A01
	UNCLASSIFIED	

EPA Form 2220-1 (9-73)

REPRODUCED BY
NATIONAL TECHNICAL
INFORMATION SERVICE
U. S. DEPARTMENT OF COMMERCE
SPRINGFIELD, VA. 22161

EPA-600/J-76-034
JOURNAL ARTICLE

a quantitative evaluation of the pollutant standards index

Wayne R. Ott and **William F. Hunt, Jr.**

U. S. Environmental Protection Agency

The U. S. Environmental Protection Agency has recommended a daily air pollution index for use by State and local air pollution control agencies—the Pollutant Standards Index (PSI). The new index makes use of "segmented linear functions" which convert measured concentrations of each air pollutant into a normalized number. The overall index is reported for the pollutant with the maximum subindex value, or the "critical pollutant."

To evaluate the index during its design phases, a computer analysis was undertaken using air quality data from selected cities. This analysis examined the statistical properties of the new index, the time series of index values, and the frequency distributions of descriptor words and critical pollutants. The index was found to be an effective tool for compactly representing a year of air quality data at any air monitoring station. It showed sufficient day-to-day variation to give useful information to the public on daily air quality levels. This paper provides sufficient technical detail for the engineer who wishes to apply the index on a large-scale computer, minicomputer, or hand calculator.

In May 1974, Thom and Ott[1] began a survey of all the air pollution indices in existence, resulting in publication of a compendium which documented the great diversity of air pollution index systems currently used in the United States and Canada. Some of the findings of this compendium were reported in a Critical Review article[2] published in the May 1976 issue of *JAPCA*. This compendium introduced a special classification coding system to compare air pollution indices, suggested the concept of a nationally standardized index, and offered a candidate structure for a uniform index—the Standardized Urban Air Quality Index (SUAQI).

As a result of this study, the Council on Environmental Quality (CEQ) established a Federal Interagency Task

Reprinted from **APCA JOURNAL**, *Vol. 26, No. 11, November 1976*

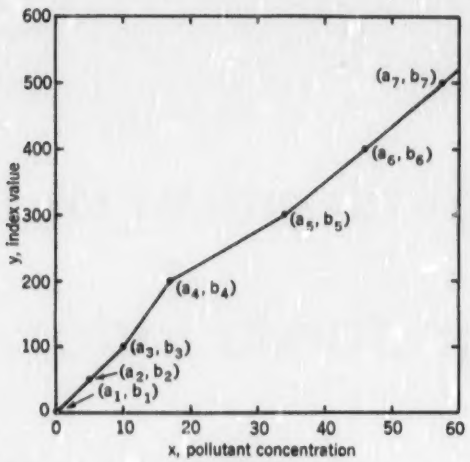


Figure 1. Example of a segmented linear function.

Force on Air Quality Indicators in July 1975, with members from the Council, the Environmental Protection Agency (EPA), the National Oceanic and Atmospheric Administration, and the National Bureau of Standards.[3] In August 1976, the Environmental Protection Agency proposed the Pollutant Standards Index (PSI) to the Task Force as a recommended index for use by State and local air pollution control agencies. The PSI index evolved from the SUAQI structure and from intermediate index structures and concepts developed along the way.[4-6] It is based on the preferred characteristics of metropolitan indices currently in use and on the health effects of various air pollutants. The PSI index, which is discussed in the News Focus section of this issue of the *APCA Journal*, was published in the *Federal Register*[7] in September and appears in a Guidelines document[8] available on request from the Environmental Protection Agency.

During the developmental phases of the PSI index, a computer study was undertaken to determine how the index performed with actual air quality data from U. S. cities. This paper presents the results of this computer evaluation. It also is intended to provide all necessary equations and coefficients that are required by the engineer wishing to program the PSI index on a computer, minicomputer, or programmable calculator.

Structure of PSI

Using the terminology developed in the recent compendium of existing air pollution indices,[1,2] PSI structurally is a "$5B_2B$" type of index. This means that the index includes up to 5 pollutants; it incorporates a "segmented linear function"; it reports on the maximum pollutant concentration ("maximum mode"); and its descriptor words relate to the National Ambient Air Quality Standards (NAAQS), the Federal Episode Criteria, and the Significant Harm Levels.

The 5 pollutants included in PSI are carbon monoxide (CO), photochemical oxidant (principally O_3), nitrogen dioxide (NO_2), sulfur dioxide (SO_2), and total suspended particulates (TSP). A sixth variable—the product $SO_2 \times$ TSP—is computed and is included in the index equation.

A segmented linear function consists of discrete coordinates (i.e., "breakpoints") joined by straight line segments (Figure 1). If the x and y coordinates of the breakpoints are represented by $(a_1, b_1), \ldots, (a_i, b_i), \ldots, (a_7, b_7)$, where x is pollutant concentration and y is the subindex value for a given variable, the following equation can be used to represent a segmented linear function:

First Segment

$$y = \frac{b_2 - b_1}{a_2 - a_1}(x - a_1) + b_1 \quad \text{for } a_1 < x \leq a_2 \tag{1}$$

Any Segment

$$y = \frac{b_{i+1} - b_i}{a_{i+1} - a_i}(x - a_i) + b_i \quad \text{for } a_i < x \leq a_{i+1} \tag{2}$$

where $i = 1, 2, \ldots, 6$.

PSI incorporates, as breakpoints, the National Ambient Air Quality Standards and the Federal Episode Criteria and Significant Harm Levels (Table I). For simplicity, we can write Eq. (3) as a product involving the coefficient α_i, the slope of the line segment, as follows:

$$y = \alpha_i (x - a_i) + b_i \quad \text{for } a_i < x \leq a_{i+1} \tag{3}$$

where $\alpha_i = \frac{b_{i+1} - b_i}{a_{i+1} - a_i}$

The coefficients for calculating the index, a_i, b_i, and α_i, are listed in tabular form in metric units (Table I) and in ppm (Table II).

The PSI index then is reported as the maximum of subindices calculated using segmented linear functions, Eq. (3), for each pollutant variable j, where y_j represents a subindex:

$$\psi = \max \{y_1, \ldots, y_j, \ldots, y_6\} \tag{4}$$

The pollutant responsible for the maximum index value—that is, the "critical

Table I. Coefficients for calculating the index in metric units.[a]

		CO		Oxidant		NO_2		SO_2		TSP		TSP × SO_2	
i	b_i	a_i	α_i	a_i	α_i	a_i	α_i	a_i	α_i	a_i	α_i	a_i	α_i
1	0	0.0	10.000000	0.0	.62500000	[b]	[b]	0.0	.62500000	0.0	.66666667	[b]	[b]
2	50	5.0	10.000000	80.0	.62500000	[b]	[b]	80.0	.17543860	75.0	.27027027	[b]	[b]
3	100	10.0	14.285714	160.0	.41666667	[b]	[b]	365.0	.22988506	260.0	.86956522	[b]	[b]
4	200	17.0	5.8823529	400.0	.25000000	1130.	.088495575	800.0	.12500000	375.0	.40000000	65000.	.000510204
5	300	34.0	8.3333333	800.0	.50000000	2260.	.135135135	1600.0	.20000000	625.0	.40000000	261000.	.000757576
6	400	46.0	8.6956522	1000.0	.50000000	3000.	.133333333	2100.0	.19230769	875.0	.80000000	393000.	.001030928
7	500	56.7	8.6956522	1200.0	.50000000	3750.	.133333333	2620.0	.19230769	1000.0	.80000000	490000.	.001030928

[a] For CO, a_i is expressed in mg/m³; for oxidant, NO_2, SO_2, and TSP, a_i is in units of μg/m³; for TSP × SO_2, a_i is expressed in μg²/m⁶.
[b] Index is not calculated; no standards or episode criteria exist at these levels.

Table II. Coefficients for calculating the index in parts per million (ppm).[a]

		CO		Oxidant		NO_2		SO_2		TSP × SO_2	
i	b_i	a_i	α_i	a_i	α_i	a_i	α_i	a_i	α_i	a_i	α_i
1	0	0.0	11.1111111	0.0	1250.00000	[b]	[b]	0.0	1666.66667	[b]	[b]
2	50	4.5	11.1111111	0.04	1250.00000	[b]	[b]	0.03	454.545456	[b]	[b]
3	100	9.0	16.6666667	0.08	833.33333	[b]	[b]	0.14	625.000000	[b]	[b]
4	200	15.0	6.6666667	0.20	500.00000	0.60	166.666667	0.30	333.333333	22.7272	1.4991837
5	300	30.0	10.0000000	0.40	1000.00000	1.20	250.000000	0.60	500.000000	91.2587	2.1666667
6	400	40.0	10.0000000	0.50	1000.00000	1.60	250.000000	0.80	500.000000	137.413	2.9484536
7	500	50.0	10.0000000	0.60	1000.00000	2.00	250.000000	1.00	500.000000	171.329	2.9484536

[a] Total suspended particulates (TSP) are not listed because they are not reported in ppm. The units of the product TSP × SO_2 is ppm-$\mu g/m^3$.
[b] Index is not calculated: no standards or episode criteria exist for these levels.

pollutant"—is identified when the index value is reported. The user may elect to report other pollutants as well, particularly if a subindex exceeds 100, the value representing a violation of the NAAQS. A typical broadcast report would be "Today's index value is 141, which means that the air is 'unhealthful.' The pollutant responsible is carbon monoxide." The descriptor words reported along with the PSI index, as it is currently designed, are: 0–50, "good"; 51–100, "moderate"; 101–200, "unhealthful"; 201–300, "very unhealthful"; 301–500, "hazardous."

Application to Data from U. S. Cities

To examine the performance of this index with actual air quality data, a FORTRAN program called INDEX. PLOT was written. This program was run on 365 days of 1974 air quality data from eight selected air monitoring stations in different cities. The index was calculated once per day, using the appropriate averaging period for each pollutant and the maximum running average concentration for the 24 hr period. While the data sets were reasonably complete, there were many instances in three cities—Chicago, Houston, and Seattle—where one or more pollutants were missing on a given day. Therefore, our discussion of the comparison of index values in different cities places emphasis on the five cities which had reasonably complete data—downtown Los Angeles, Lennox, CA, Anaheim, CA, Camden, NJ, and Newark, NJ.

Statistical Findings

The difference between the highest and lowest PSI values (the range) did not vary greatly from city to city; the smallest range was 152.2 in Seattle and the largest range was 286.7 in Lennox (Table III). For these eight data sets, the annual arithmetic mean of PSI ranged from 44.0 in Seattle to 119.2 in downtown Los Angeles, with the standard deviations ranging from 21.3 in Seattle to 50.1 in downtown Los Angeles. The coefficient of variation (ratio of the standard deviation to the mean) remained somewhat more constant, ranging from 0.421 in downtown Los Angeles to 0.677 in Houston. This finding suggests that the "spread" of the PSI frequency distribution, as measured by its standard deviation, may be, to some degree, a function of its mean. The calculated coefficient of skewness $\sqrt{b_1}$ was always positive, ranging from 0.02 in Chicago to 2.18 in Camden. The small $\sqrt{b_1}$ in Chicago suggests a nearly symmetrical distribution of observed index values, while the distributions for the other cities have a significant right "tail"—that is, they are skewed to the right. The coefficient of kurtosis, which gives a measure of the "peakedness" of the shape of the distribution (relationship of height to width), ranges from 2.32 in downtown Los Angeles to 12.75 in Camden, while Chicago has a value of 3.08. For the normal distribution, the theoretical values for $\sqrt{b_1}$ and b_2 are 0.0 and 3.0, respectively, suggesting that the index values for Chicago tend more toward a normal distribution than a lognormal distribution.

A 3 month segment of a time series plot for one of the cities is shown in Figure 2. Each vertical line represents the index value for a day of the year, and the small letter at the top of each line denotes the critical pollutant (the pollutant with maximum subindex). The index shows considerable day-to-day variation. In Los Angeles, the predominant pollutants are CO and oxidant, with CO dominating the index during the period from mid-October to March and oxidant the remaining months. While violations of the NAAQS (values of PSI > 100) occur throughout the year, the highest values of PSI occurred in December.

Comparison of Cities

The cumulative frequencies for the index data can readily be plotted on logarithmic probability paper (Figure 3) to give a visual representation of the results for each city. On this paper, lack of curvature, or "straightness" of the plots usually is interpreted to mean that the distribution tends toward log-normality. These lines show the difficulty of ranking the severity of pollution levels in the five cities just in terms of mean values. If ranked from most severe to least severe using the arithmetic means, the highest index values would be in downtown Los Angeles, followed by Lennox, Anaheim, Newark, and Camden. If, however, rankings are based on the highest index values observed (for example, values greater than $\psi = 200$), Lennox and downtown Los Angeles are

Table III. Statistical characteristics of the PSI index when applied to 8 U. S. cities.

City	n	$\bar{x}$	s	cov	$\sqrt{b_1}$	b_2	Lowest	Highest	Range
L.A. Downtown	364	119.2	50.1	.421	.37	2.32	25.0	257.1	232.1
Lennox, CA	365	112.8	71.2	.631	.86	2.43	20.0	306.7	286.7
Anaheim, CA	365	81.5	43.0	.528	.97	3.73	12.5	255.0	242.5
Chicago, IL	364	98.5	42.0	.426	.02	3.08	6.3	224.1	217.9
Newark, NJ	353	65.5	37.6	.575	1.34	5.29	0.0[a]	212.4	212.4
Camden, NJ	356	50.7	28.2	.556	2.18	12.75	0.0[a]	232.4	232.4
Houston, TX	363	59.3	40.2	.677	.83	3.75	0.0[a]	208.2	208.2
Seattle, WA	357	44.0	21.3	.485	1.32	6.68	0.0[a]	152.2	152.2

Explanation of terms:
n = number of days for which data were available.
$\bar{x}$ = arithmetic mean
s = arithmetic standard deviation
cov = coefficient of variation, $s/\bar{x}$
$\sqrt{b_1}$ = coefficient of skewness
b_2 = coefficient of kurtosis
[a] In these computer runs, the index was set to zero if data for all five air pollutants were missing; thus the lowest value is underestimated.

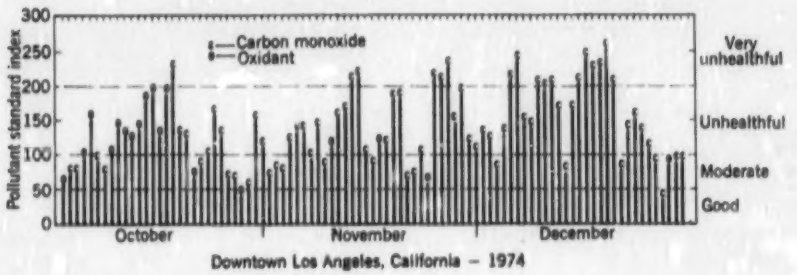


Figure 2. Example of time series plot of PSI values.

reversed. These two different orderings of the data sets illustrate the difficulty of attempting to compare air quality levels in different cities using this index. These inconsistencies are not unique to the PSI index and will arise for most other index structures as well.

Another serious difficulty in comparing cities stems from nonuniformity in quality control practices at particular air monitoring stations and from nonstandardization of station sites. As reported elsewhere in the literature,[9,10] variation in monitoring station sites tends to bias measured concentrations, particularly for CO, making it difficult to generalize about the representativeness of a given station or readily to compare cities using the data from these stations.

It is possible on one figure to plot the distributions of descriptor words for all five cities (Figure 4). These distributions differ strikingly. In Anaheim (top, left), the 365 days of the year were divided almost equally among three descriptor words: "good," "moderate," and "unhealthful," with PSI above 200 only 1% of the time. In Lennox, PSI was above 200 more often, causing 20% of the days to be designated, "unhealthful." Downtown Los Angeles gave a pattern that was different than either Anaheim or Lennox: the majority of the descriptor words (53%) fell into a single category, "very unhealthful." One of the five cities, Camden, had the lowest concentrations, with 52% of the days reported as "good," and only 5% above the NAAQS (PSI > 100). In Newark, the majority of the descriptor words were in the "good" or "moderate" ranges, and just 16% fell into the range above the NAAQS. By comparison, 62% were in the range above the NAAQS in Los Angeles. For all five cities, values of PSI above 300 were very rare: there was only one instance of a "hazardous" descriptor word (in Lennox).

The distributions of critical pollutants (pollutants with the highest subindex value) also showed considerable variation from city to city, as shown by the histograms on the right half of Figure 4.

In the California cities, the most common critical pollutants were CO and oxidant, both of which are related to the motor vehicle. Taken together, they accounted for 89% of the days in Lennox, 92% of the days in Anaheim, and 99% of the days in downtown Los Angeles. Although CO and oxidant were the most frequently occurring critical pollutants, the proportion of CO days relative to oxidant days varied from city to city. In Anaheim, oxidant was the critical pollutant for 71% of the days. In Lennox, the situation was almost reversed: CO showed up as the critical pollutant for 75% of the days, and oxidant showed up only 14% of the time. In the three California cities, SO_2 accounted for a very small proportion of the days—never more than 2%—while TSP varied from 1% in downtown Los Angeles to 8% in Lennox and Anaheim.

In the two New Jersey cities, Newark and Camden, SO_2 showed up as the critical pollutant more frequently than in the California cities. In Camden, it was the most frequently occurring critical pollutant, with 39% of the days, while in Newark it was the second most frequent critical pollutant, with 11% of the days. Although SO_2 showed up more frequently in the two New Jersey cities than in the three California cities, CO and oxidant together still accounted for more than half of the days—58% in Camden and 87% in Newark. In the two eastern cities, TSP accounted for a very small share of the total critical pollutant days—3% in Camden and 2% in Newark. The low incidence of TSP days in all cities is due, in part, to the smaller amount of TSP data available from high-volume sampler measurements. The product $SO_2 \times TSP$ never appeared as the critical pollutant, nor did NO_2.

In summary, the highlights of applying PSI, as currently structured, to data from these five cities are as follows:

- The dominant pollutants were CO and oxidant, together accounting for more than 89% of the days in the California cities and over 58% of the days in the New Jersey cities.
- Sometimes CO occurred more frequently than oxidant; sometimes the reverse occurred.
- In addition to CO and oxidant, SO_2 was a common critical pollutant in the New Jersey cities, alone accounting for 39% of the days in Camden and 11% of the days in Newark.
- NO_2 never appeared as the critical pollutant.
- $SO_2 \times TSP$ never appeared as the critical pollutant.
- The most severe descriptor word, "hazardous," occurred only once.
- The ranking of cities is a complex

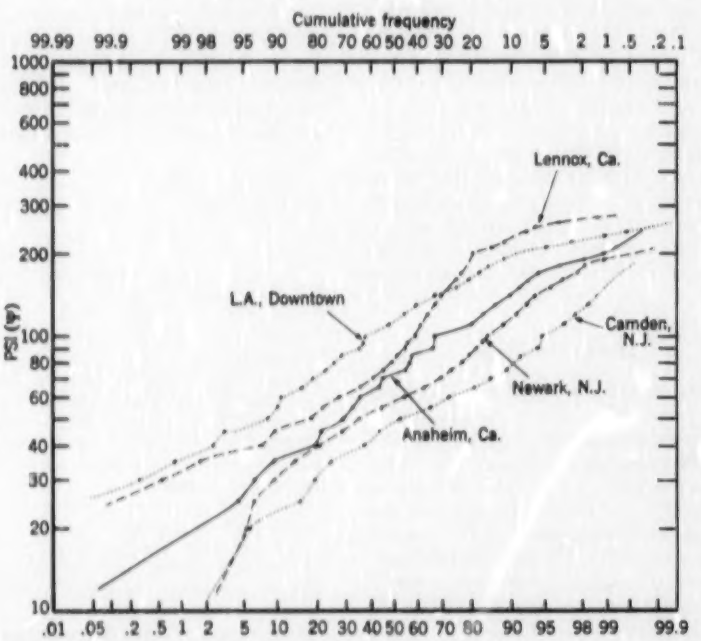
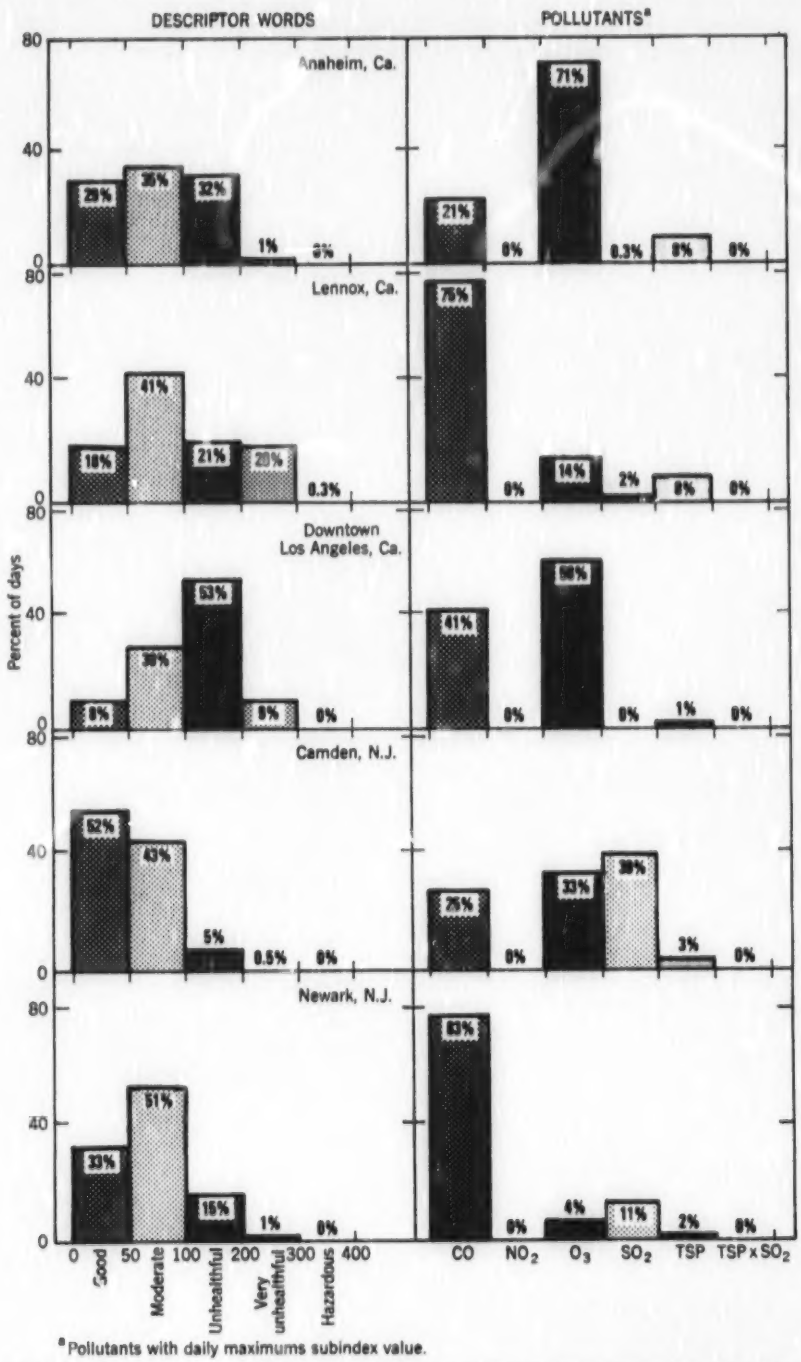


Figure 3. Logarithmic probability plots of PSI for five U. S. cities.



[a] Pollutants with daily maximums subindex value.

Figure 4. Distributions of PSI descriptor words and pollutants in the three California and two New Jersey cities for 1974.

task, not easily accomplished with this or with other indices.

Conclusion

Our evaluation of the PSI air pollution index recommended by EPA suggests that the index reflects day-to-day variations of air pollutant concentrations in a reasonable fashion. The index appears to be a useful tool for presenting and interpreting air quality data. It also can be used in trend studies to display annual air quality levels at monitoring stations in U. S. cities. Its mathematical properties and the simplicity of the data summaries it generates, as revealed by our computer programs, make the index an effective means for compactly representing a year of air quality data for any monitoring station or city. Although these displays permit different cities to be compared in a very approximate fashion, such comparisons are complex and must be carried out with great caution due to variations in air monitoring practices.

References

1. G. C. Thom and W. R. Ott, *Air Pollution Indices: A Compendium and Assessment of Indices Used in the United States and Canada*, Ann Arbor Science Publishers, Ann Arbor, MI, 1976.
2. W. R. Ott and G. C. Thom, "A critical review of air pollution index systems in the United States and Canada," *J. Air Poll. Control Assoc.* **26:** 460 (1976).
3. G. C. Thom and W. R. Ott, "Progress toward a uniform air pollution index," *J. Air Poll. Control Assoc.* **25:** 1157 (1975).
4. W. F. Hunt, W. Cox, W. R. Ott, and G. C. Thom, "A Common Air Quality Reporting Format: Precursor to an Air Quality Index," *Proceedings of the Fifth Annual Environmental Engineering and Science Conference*, Louisville, KY, March 3–4, 1975, pp. 101–121.
5. G. C. Thom and W. R. Ott, "A proposed uniform air pollution index," *Atmos. Environ.* **10:** 261 (1976).
6. G. C. Thom, W. R. Ott, W. F. Hunt, and J. B. Moran, "A Recommended Standard Air Pollution Index," presented at the 171*st* National Meeting of the American Chemical Society, New York, NY, Apr. 7, 1976.
7. Environmental Protection Agency, "Guideline for public reporting of daily air quality," *Federal Register* 41 (174): 37660 (Sept. 7, 1976).
8. Environmental Protection Agency, "Guideline for Public Reporting of Daily Air Quality—Pollutant Standards Index (PSI)," Office of Air Quality Planning and Standards, Publication No. EPA-450/2-76-013, Research Triangle Park, N. C. 27711, Aug. 1976.
9. W. R. Ott, "Development of Criteria for Siting Air Monitoring Stations," Paper No. 75-14.2, presented at the 68*th* Annual Meeting of the Air Pollution Control Association, Boston, June 1975.
10. F. L. Ludwig and J. H. S. Kealoah, "Selecting Sites for Carbon Monoxide Monitoring," prepared by Stanford Research Institute for the Environmental Protection Agency, Publication No. EPA-450/3-75-077, Research Triangle Park, NC 27711, Sept. 1975.

Dr. Ott is Senior Systems Analyst for Monitoring Quality Assurance, Office of Monitoring and Technical Support, Office of Research and Development, U. S. Environmental Protection Agency, 401 M. Street, S.W., Washington, DC 20460. Mr. Hunt is Chief, Data Analysis Section, Monitoring and Data Reports Branch, Office of Air Quality Planning and Standards, U. S. Environmental Protection Agency, Research Triangle Park, NC 27711.

This is a revised version of Paper No. 76-29.4 which was presented at the 69th Annual Meeting of APCA at Portland in June 1976.